Rise in antisemitism on both mainstream and fringe social media platforms following Hamas' terrorist attack

Updated 7 February 2024 to include longer term alt-tech platform data, originally published 31 October 2023.

By Hannah Rose (https://www.isdglobal.org/isd_team/hannah-rose/), Jakob Guhl (https://www.isdglobal.org/isd_team/jakob-guhl/) and Milo Comerford (https://www.isdglobal.org/isd_team/milo-comerford/)

ISD analysis – rooted in a bespoke hate speech classifier – finds an over 50-fold increase in the absolute volume of antisemitic comments on YouTube videos about the Israel/Palestine conflict, following Hamas' attacks. The data also demonstrated a 2.4 times increase in the overall proportion of antisemitic messages during the same period. Cross-platform monitoring has further shown increased online threats against Jewish communities across both mainstream and fringe social media.

Hamas' terrorist attack on Israel on 7 October, followed by the Israel-Hamas war, has led to a significant increase (https://www.report-antisemitism.de/bundesverband-rias/) in antisemitism targeting communities. Sadly, this pattern is not new, where antisemitism against Jewish communities outside of Israel is known to increase (https://www.jpr.org.uk/reports/how-conflict-israel-and-gaza-affects-jews-uk-and-what-you-can-do-about-it) following heightened tensions in Israel and Palestine.

This research brief examines the rapid increase of online antisemitism on YouTube and alternative social media platforms such as Telegram and 4chan immediately after the terrorist attack.

In the 17 days after the Hamas terror attack, the UK's Community Security Trust (CST), a charity monitoring antisemitism, recorded 600 antisemitic incidents (https://cst.org.uk/news/blog/2023/10/23/antisemitic-incidents-23-october-update). This marks a 641% increase from the same period in 2022. Of these incidents, 185 occurred online, including abusive language, harassment, and threats of violence. Similarly, in the week following the attack, German monitoring organisation RIAS recorded 202 incidents (https://www.report-antisemitism.de/about/), a 240% increase on the same period the previous year. Additionally, hate crime data (https://news.sky.com/story/antisemitic-and-islamophobic-hate-crimes-up-amid-israel-

hamas-war-according-to-metropolitan-police-12988310) from London's Metropolitan Police, recorded an astonishing 1,353% rise in antisemitic offences between 1 and 18 October.

This sits in the backdrop of a broader rise of antisemitism across both extremist and mainstream constituencies. Antisemitic conspiracy theories, particularly in the online space, rose in response to the COVID pandemic, with ISD data (https://www.isdglobal.org/isd-publications/the-rise-of-antisemitism-during-the-pandemic/) demonstrating a seven and thirteenfold rise in French and German language posts on Twitter, Facebook and Telegram. ISD and CASM research (https://www.isdglobal.org/isd-publications/antisemitism-on-twitter-before-and-after-elon-musks-acquisition/) also found a doubling in antisemitism on Twitter/X after Elon Musk's takeover.

Antisemitism on social media poses a challenge for analysts due to its implicit nature, often relying on context to determine whether a statement is antisemitic. Because of this, keyword-based approaches for surfacing messages will often produce incomplete sets. This research relies instead on automated detection models and builds on a machine learning classifier that was previously created to measure the level of antisemitism on X (https://www.isdglobal.org/isd-publications/antisemitism-on-twitter-before-and-after-elon-musks-acquisition/)(formerly Twitter) for assessing antisemitism in YouTube comments.

This study also utilises keyword lists of antisemitic slurs to estimate the level of antisemitism on alternative social media platforms, and discusses some of the emergent key trends, including threats to Jewish public figures and responses by extremist communities. YouTube was chosen for analysis due to similar content volume and message length to X, meaning that we could adjust a classifier from previous research (https://www.isdglobal.org/wp-content/uploads/2023/03/Antisemitism-on-Twitter-Before-and-After-Elon-Musks-Acquisition.pdf). Furthermore, the emergence of the EU's digital regulation framework, the Digital Services Act, prompted an analysis of antisemitism on a Very Large Online Platform (VLOP) to better understand and evaluate platform responsibilities. It would have been ideal to conduct comparative analysis across other platforms such as Meta-owned platforms, X, and TikTok, which could be conducted with the training of platform-specific classifiers over a longer period of time. However, current data restrictions imposed by those platforms prevent systematic and comparative analysis by most research organisations.

Antisemitism on YouTube

Method

Distinguishing between legitimate criticism of Israel and antisemitism can be challenging, particularly when specific keywords are absent. Therefore, relying purely on keyword searches on mainstream social media will often fail to capture the nuance between the two, especially when specific keywords are not used. To address this, our research uses a classifier-based approach to detect antisemitic comments, based on previous work by ISD and CASM Technology measuring hate speech in the UK (https://www.ofcom.org.uk/research-and-data/online-research/Online-terrorism-violence-and-hate) and antisemitism on X (https://www.isdglobal.org/isd-publications/antisemitism-on-twitter-before-and-after-elon-musks-acquisition/) (formerly Twitter).

Our methodology (outlined in full in an annex) involved collecting comments from videos related to Israel and Gaza between 3 to 13 October, to then compare to the period before the attacks. Over 5 million comments from 11,000 videos were gathered. These comments were then filtered using keywords list, combination logic and language filters resulting in a subset of over 45,000 comments potentially related to English-language antisemitism. These posts were then run through an 'ensemble' of existing hate speech classifiers and keyword annotators, with a final layer comparing the performance of these annotations

against a model trained by ISD subject-matter experts on almost 2,000 manually labelled posts. Posts were coded in line with the working definition of antisemitism of the International Holocaust Remembrance Alliance (IHRA) (https://www.holocaustremembrance.com/resources/working-definitions-charters/working-definition-antisemitism). Coders tended towards a more conservative interpretation of the working definition to ensure a high-certainty threshold for inclusion. This final model demonstrated an 82% overall accuracy in identifying antisemitic posts.

Findings

In the three days after Hamas' attack, the absolute number of antisemitic comments on conflict-related YouTube videos increased by 4963%, when compared to the previous three days. This increase continued after the first three days after the attack, with an additional 37.2% increase in the number of antisemitic comments from 10-12 October compared to 7-9 October.

While the increase in absolute numbers of comments is also partially explained by an increase in the number of videos focusing on Israel/Gaza, the proportion of antisemitic comments has also increased by 242% when comparing the daily average before and after the attack.

A total of 15,720 antisemitic YouTube comments were identified after the attack. This increase coincided with a peak in overall volume of collected videos, as well as Israel's retaliatory effort, including the mobilisation of troops and the targeting of sites in Gaza.

Made with Flourish (https://public.flourish.studio/visualisation/15506689/?utm_source=embed&utm_campaign=visualisation/15506689)

Graph 1: Number of antisemitic YouTube comments by day before and after the Hamas attacks

While antisemitic comments began sharply rising immediately after the October 7 terror attacks, the highest number of such comments was recorded on October 10. Antisemitic comments on YouTube remained consistently high in the week following Hamas' attack, demonstrating the sustained increase in hostility towards Jews on social media, extending beyond the initial attack. Such findings on the increase of antisemitism on social media mirror data recently published by Decoding Antisemitism, (https://decoding-antisemitism.eu/publications/celebrating-terror-antisemitism-online-after-the-hamas-attacks-on-israel/) evidencing a 54.7% increase in antisemitism UK YouTube comments.

Among the significant volume of identified antisemitic comments, a number of broad themes clearly aligned with the IHRA's working definition of antisemitism. This included dehumanising language, often used with reference to Jewish people, including slur words or comparing Israelis or Zionists to Nazis, and their actions to those of Hitler. Additionally, common conspiracy theories were integrated into the context of the conflict, including that Jewish people are 'Khazars' (https://www.isdglobal.org/digital_dispatches/an-antisemitic-conspiracy-theory-is-being-shared-on-telegram-to-justify-russias-invasion-of-ukraine/) or not descendants of Israelites and thereby 'fake Jews'. This fed into the narrative that Jews are a European 'problem' forced onto Palestinians in the aftermath of the Holocaust, both denying Jewish rights to self-determination and re-surfacing 'othering' narratives of Jews. Existing conspiracy theories about Jewish control of political, media and financial institutions were applied to the conflict, commonly around Israel's perceived control of the US government's political and financial response. The deicide myth (that Jews killed Jesus) was used to justify and celebrate Jewish deaths, alongside Black Hebrew Israelite (https://www.adl.org/resources/backgrounder/extremist-sects-within-black-hebrew-israelite-movement) language such as the 'Synagogue of Satan' (https://extremismterms.adl.org/glossary/synagogue-satan). Some antisemitic comments built on 9/11 conspiracy theories suggesting that Jewish people or Israel committed the terrorist attacks in order to justify wars in Afghanistan in Iraq, claiming that Hamas' attack was a 'false

flag (https://www.adl.org/glossary/false-flag)' performed by Israel for ulterior motives which might include political power, financial benefit or to lay the groundwork for an offensive in Gaza.

During comment annotation, ISD identified a range of material that fell into the grey area between legitimate criticism of Israel and antisemitism. For example, comments describing Israel as "Israhell" or accusing it of apartheid may be offensive to many but do not necessarily fall under the concrete examples of antisemitism provided by the IHRA (e.g., "denying the Jewish people their right to self-determination") unless there was further context. While both terms suggest fierce criticism of Israeli policies and behaviour towards the Palestinians, neither automatically implies that "the existence of a State of Israel is a racist endeavour" (another IHRA example). Similarly, the protest slogan "from the river to sea, Palestine will be free" (https://www.ajc.org/translatehate/From-the-River-to-the-Sea) was not automatically coded as antisemitic due to multiple possible interpretations. While one interpretation of the slogan is that it is a genocidal call for the destruction of Israel, another is a call for a bilateral state solution, and without context this phrase cannot be definitively identified as antisemitic (https://news.met.police.uk/news/update-met-response-to-terror-attacks-in-israel-and-ongoing-military-action-in-gaza-474080). Given the validity of different interpretations and understandings of antisemitism, the above data may underestimate the level of antisemitism in YouTube comments.

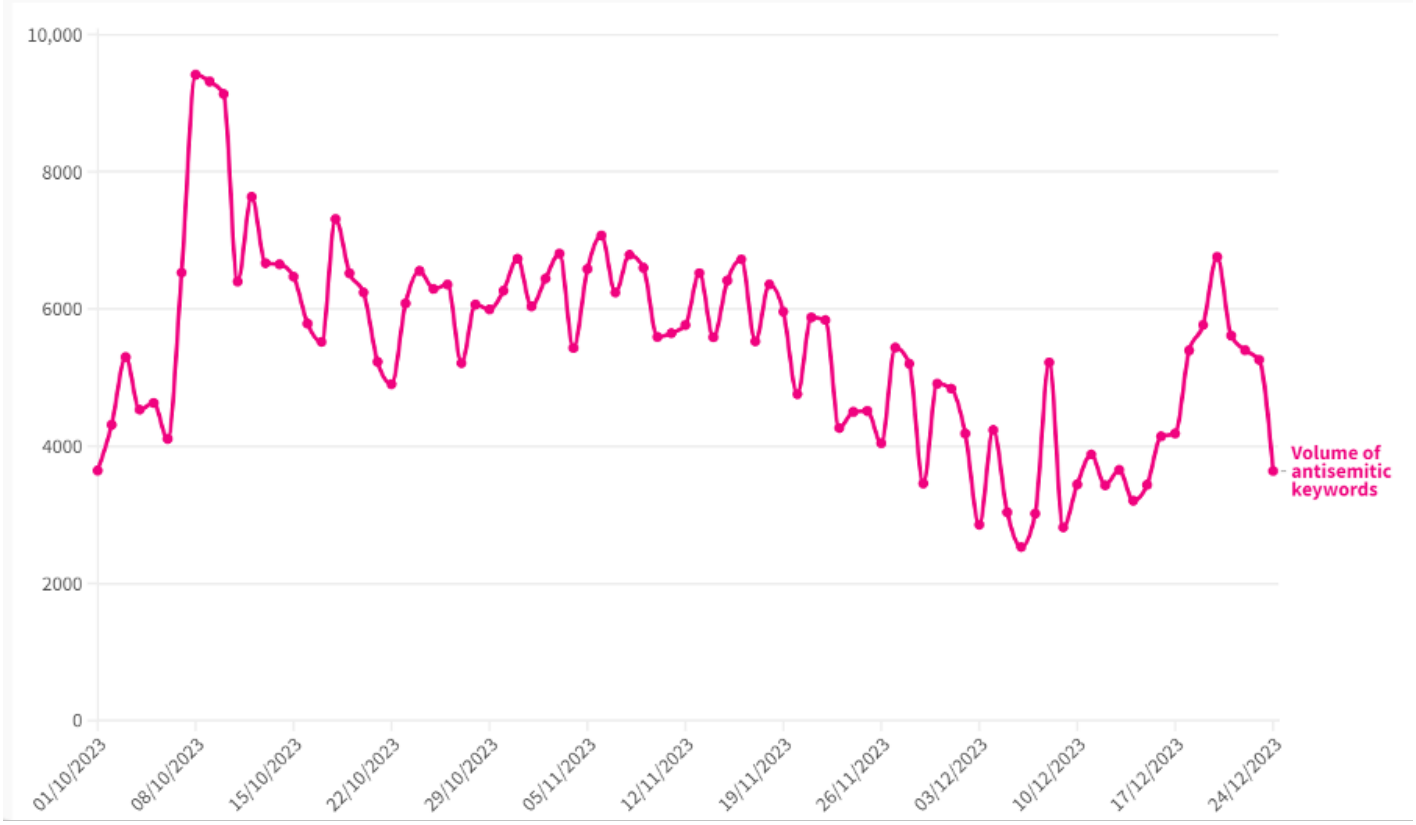
Antisemitism on alt-tech platforms

Made with Flourish (https://public.flourish.studio/visualisation/15410358/?utm_source=embed&utm_campaign=visualisation/15410358)

Graph 2: Posts showing antisemitic keywords on different alternative social media platforms

In addition to analysing YouTube comments, ISD sought to understand mobilisation on fringe platforms known for more permissive environments for hate and extremism. In cases where an existing classifier did not exist for the platforms and API access was limited, keyword-based approaches can be deployed for the analysis of overt antisemitic content. For this analysis, a list of keywords which have a high likelihood of antisemitism was used, including slurs and terms used by extreme-right ecosystems. Given the presence of extremist ecosystems (https://cst.org.uk/news/blog/2020/06/11/hate-fuel-the-hidden-online-world-fuelling-far-right-terror) on fringe social media platforms, more overt expressions of antisemitism are commonly found, therefore rendering a keyword approach useful for this context. The keyword list was entered into a tool for measuring the frequency of keywords across a range of fringe social media platforms. This provided a snapshot of the level of antisemitic content, omitting posts which may have been antisemitic but did not include a keyword. Not every channel or forum on every platform was surveyed by the tool, and therefore keyword frequency on platforms such as Telegram, where individual channels had to be added, is likely to also be an underestimate.

The above graph visualises a spike in posts containing antisemitic keywords on platforms including 4chan, Bitchute, Gab and Telegram, among others, in the immediate aftermath of the October 7 Hamas terrorist attack. The three days following the attack saw over a three-fold increase in volume of antisemitic keywords on the previous three days, followed by a sustained increase over the next week as the conflict escalated and Israel's responded militarily. Due to the time-limited nature of data collection and the high volume of posts, longer-term analysis is unavailable. The data demonstrates a noted high volume of antisemitic content on 4chan, mirroring that 4chan provides anonymous users with space to share explicit and often violent antisemitic narratives and imagery. The themes and targets of such antisemitic material will be explored below.



Graph 3: antisemitic keyword volumes on fringe platforms, 1 October to 24 December 2023

In order to understand the longer-term impact of October 7 on antisemitism on fringe platforms, analysts collected keyword volumes until 24 December 2023. In this two-and-a-half-month period, over 460,000 posts were identified as using antisemitic keywords and phrases on fringe platforms. The daily average of posts containing antisemitic keywords until 24 December rose 25% after the October 7 attack. During the month of October, there was a 50% increase in the daily average of antisemitic language after October 7. This evidences both the length of the initial shock of October 7, and its sustained impact on the level of antisemitic abuse and incitement towards Jewish communities on alternative social media platforms.

Threats to Jewish figures and officials

Threats and violent rhetoric aimed at Jewish communities have been evidenced since before the recent conflict. However, the Israel-Gaza war heightened the threat landscape, evidenced by a surge in antisemitic language on social media platforms, resulting in additional threats towards Jewish communities. These threats took various forms online through graphic imagery or direct calls for violence, and in some cases can manifest in offline harms.

Online antisemitic ecosystems on alternative social media platforms have been used by extremist actors to share tactics and call for violence against Jewish people, both in Western countries and in Israel. On Telegram, for example, graphic images of deceased Israelis have been used to call for violence against Jewish people, which ISD has decided not to include in this briefing due to their graphic nature. Jewish public officials, including elections officials, have been threatened on- and offline. Jewish sites, such as synagogues, have been a common target of antisemitic violence, with the vandalism of a synagogue in Porto (https://www.bing.com/ck/a?!&&p=7c25d552c58b8214JmltdHM9MTY5ODEwNTYwMCZpZ3VpZD0yZjRkMmM0MC1mN2M2LTZhNDEtMThkOS f7c6-6a41-18d9-3f7ef6e16bd4&psq=synagogue+graffiti+porto&u=a1aHR0cHM6Ly93d3cudGltZXNvZmlzcmFlbC5jb20vcG9ydG8

the firebombing of a Berlin synagogue, (https://www.bbc.co.uk/news/world-europe-67145473)the torching of the historic synagogue in Tunisia (https://www.thejc.com/news/world/historic-tunisian-synagogue-set-on-fire-by-pro-palestine-rioters-2GRLnK6GlSMpvPWuSXunZz), and in US synagogues (https://www.necn.com/news/local/synagogues-in-southern-mass-and-r-i-receive-bomb-threats/3067407/).

Antisemitism among extremist communities

Extremist communities are present on mainstream and fringe social media platforms. Islamist groups, specifically, such as Al-Qaeda have published content on Facebook encouraging terrorist attacks against Jewish targets in Israel and Israeli embassies abroad. They framed Arab states and leaders as servants of Israel and the US, calling for their assassination and advancing the narrative of a global Zionist conspiracy influencing world events. Islamic State published an article in their weekly newsletter titled 'Practical Steps to Fight the Jews' urging attacks against Jews in the US and Europe, as well as in Israel, concluding with 'the Jews have yet to live through the Holocaust.'

In their responses to the events in Israel and Gaza, Islamist groups avoid using the word 'Israel' and instead refer to 'the enemy', 'the occupation', 'the Zionists', or just 'the Jews', the latter two of which can assign responsibility for the actions of the state of Israel to Jews worldwide. The groups often emphasise that the so-called "enemy" indiscriminately kills women, children and the elderly. Content praising the attack, Hamas, the Al-Qassem Brigades and Palestinian Islamic Jihad (PIJ) has been widespread on social media, including 160 branded posts from the terrorist organisations which were viewed over 16 million times on X alone (https://www.isdglobal.org/digital_dispatches/illegal-violent-terrorist-content-relating-to-hamas-israel-conflict-reaches-millions-on-x/). The content in some cases includes the official name given to the attack by Al-Qassem Brigades, 'Al-Aqsa Storm' as a hashtag, and originates from Telegram channels. 'Resistance Axis' pro-Iran accounts similarly celebrated the attack and excused the targeting of Israeli civilians.

Telegram groups, where extremist movements commonly converge in a network known as 'Terrorgram' (https://hopenothate.org.uk/chapter/the-terrorgram-network-a-spiral-towards-bloodshed/), have praised the Hamas attacks and called for further violence against Jews. Accelerationist (https://www.isdglobal.org/explainers/accelerationism/) networks support Hamas' perceived goal of a Jewish genocide and support any inter-group violence which they see as hastening the collapse of society. It is not uncommon to see cross-posting of operational and even ideological material between Islamist and extreme-right online ecosystems. For example, extreme-right networks have previously shown solidarity with Palestinian militants against the perceived, common 'one struggle' against Jewish people. The below image, pulled from an extreme-right Telegram channel during the last Israel-Gaza war in May 2021, is intended to depict a Palestinian terrorist and a neo-Nazi united by their common anti-Jewish 'struggle'.

Figure 1: Meme showing commonality between neo-Nazis and Salafi-jihadis

Neo-Nazi accelerationist channels particularly pointed out the guerilla-style tactics that Hamas employed and how the lessons learned from the attacks ought to be applied in the US. Neo-Nazi accelerationists also shared operational advice on how to attack critical infrastructure in Israel such as electrical substations to further destabilise the country. This advice resembles the frequent calls for attacks of critical infrastructure (https://gnet-research.org/2022/09/12/analysing-terrorgram-publications-a-new-digital-zine/) in the US.

Additionally, radicalised conspiracy movements were quick to incorporate tropes and common narratives into the conflict, in some cases framing the attacks into existing conspiracy frameworks. For example, posts by QAnon-affiliated Telegram pages and conspiratorial news sites offered accounts that Israel must have known about the attack in advance and allowed it to happen, drawing comparisons to the September 11 attacks or Pearl Harbor.

Conclusions

This research briefing highlights the surge in volume and nature of online antisemitism following Hamas' attack on Israel. It reveals spikes in antisemitic abuse and conspiracy theories, along with a sustained increase as Israel mobilised against Hamas. Thematic analysis of YouTube comments has given insights into how antisemitic conspiracy theories are evolving in the context of Hamas' terrorist attack and the subsequent war. Understanding the spreading of antisemitic narratives online can be used to protect Jewish communities from harms online and their offline repercussions through informing campaigns and security efforts.

Antisemitic content, especially when used to target Jewish individuals and institutions and incite violence, can be illegal. Where this is the case, platforms must urgently address the spreading of this content under both digital regulation frameworks and their own terms of service. However, much of the antisemitic content identified on YouTube was 'borderline' in nature or in some legal jurisdictions, legal. Platforms in these cases must go beyond merely their legal duties to ensure that harmful conspiracies and hateful language is not being spread. Given the current polarisation and heightened threat landscape, social media companies must ensure that the content spread on their platforms does not contribute to the rising threat to Jewish communities globally.

***

Methodological annex

For the analysis, videos were collected using the YouTube API from 4 to 12 October via a list of keywords relevant to the current conflict in Israel and Palestine, and comments from the relevant videos were collected. In total, 5,070,851 comments were collected from 11,220 videos. Comments were filtered using a second keyword list of terms with a likelihood of relevance to English-language antisemitism, and combination logic. This ensured that that nearly all comments were in English while including words that are not formally English but could be understand by English speakers (e.g. Israhell). This resulted in a dataset of 45,375 comments which were run through existing annotators. A sub-set of 1,970 comments from this dataset was then manually annotated to establish whether comments were relevant to antisemitism. This step was taken to aid an antisemitism classifier designed to detect antisemitic posts. Posts were coded in line with the working definition of antisemitism of the International Holocaust Remembrance Alliance (IHRA) (https://www.holocaustremembrance.com/resources/working-definitions-charters/working-definition-antisemitism). Coders tended towards a more conservative interpretation of the working definition to ensure a high-certainty threshold for inclusion.

The training data was split from the dataset and used to train separate annotators but was not used to train or evaluate XGBoost subsequently. Each post was then run through an ensemble classifier consisting of 57 annotations, including 18 publicly available machine learning models [1], 7 exact wordlists, 3 sub matching word lists, 3 models trained for this project (2 from roberta-base and one using sentencebert) and 4 models taken from previous projects analysing hate speech in the UK and antisemitism on Twitter. The final layer of the ensemble classifier was an XGBoost algorithm which takes these annotations as an input along with 1,199 posts labelled by ISD analysts (278 of which were labelled as antisemitic) and 300 test posts (102 of which were antisemitic). This produced the final model with a class one precision of 0.72 recall of 0.77, and an overall F1 score of 0.82.

The classifier model used to detect antisemitic posts for this briefing was based on previous work by ISD and CASM measuring hate speech in the UK (https://www.ofcom.org.uk/research-and-data/online-research/Online-terrorism-violence-and-hate) and antisemitism on X (https://www.isdglobal.org/isd-publications/antisemitism-on-twitter-before-and-after-elon-musks-acquisition/) (formerly Twitter). It would have been ideal to conduct comparative analysis of antisemitism across other platforms such as X. However, current restrictions to data access (https://www.wired.co.uk/article/twitter-data-api-prices-out-nearly-everyone) put in place by X means that it is not possible for the overwhelming majority of research organisations to conduct systematic analysis of antisemitism of the platform.

With thanks to Guy Fiennes, Michel Seibriger and Zoe Manzi for coding support and contributions to analysis, Francesca Arcostanzo (https://www.isdglobal.org/isd_team/francesca-arcostanzo/), Shaun Ring (CASM) and Kieran Young (CASM) for data collection and analysis.

Footnotes

1. Detoxify-multilingual, Detoxify-original, Detoxify-unbiased, IMSyPP, Dehatebert-mono, 'english-abusive-MuRIL', 'facebook-roberta-hate-speech', 'hateBERT-abuseval', 'hateBERT-hateval', 'hateBERT-offenseval', 'hatespeech-refugees', 'hatexplain', 'hatexplain-rationale-2', 'toxic-comment-model', 'twitter-roberta-base-hate', 'twitter-roberta-base-multiclass', 'Perspective', 'toxigen.general/pred_labels'

Related

(https://www.isdglobal.org/digital_dispatches/delegitimisation-of-the-demos-gegen-rechts-by-pro-russian-and-far-right-actors-in-germany/)

Delegitimisation of the 'Demos gegen Rechts' by pro-Russian and far-right actors in Germany (https://www.isdglobal.org/digital_dispatches/delegitimisation-of-the-demos-gegen-rechts-by-pro-russian-and-far-right-actors-in-germany/)

Protests in Germany against far-right extremism & anti-democratic ambitions received widespread support earlier this year, but far-right alt media & Telegram channels are using disinfo and conspiracies to say otherwise.

(https://www.isdglobal.org/digital_dispatches/the-eu-ai-act-insights-from-the-worlds-first-comprehensive-ai-law/)

The EU AI Act: Insights from the world's first comprehensive AI law (https://www.isdglobal.org/digital_dispatches/the-eu-ai-act-insights-from-the-worlds-first-comprehensive-ai-law/)

The EU AI Act (AIA) is the world's first comprehensive AI regulation that upholds democratic norms and human rights, ensures tech accountability, and prevents the spread of mis- and disinformation and its misuse by bad actors. However, some uncertainties around its implementation and enforcement remain.

(https://www.isdglobal.org/digital_dispatches/dont-you-want-to-go-on-vacation-hawking-hijrah-to-the-african-affiliates-of-the-islamic-state/)

"Don't you want to go on vacation?": Hawking 'hijrah' to the African affiliates of the Islamic State (https://www.isdglobal.org/digital_dispatches/dont-you-want-to-go-on-vacation-hawking-hijrah-to-the-african-affiliates-of-the-islamic-state/)

ISD analyses the Facebook activity of 10 Islamic State alternative news outlets, their promotion of hijrah (immigration) to Africa and the welcomed efforts to sell the continent to a base accustomed to supporting the group's actions in Syria and Iraq.

(https://www.facebook.com/ISDglobal) (https://twitter.com/ISDglobal) (https://www.linkedin.com/company/isdglobal/)

ISD

Amman | Berlin | London | Paris | Washington DC

Copyright © Institute for Strategic Dialogue (2024). Institute for Strategic Dialogue (ISD) is a company limited by guarantee, registered office address PO Box 75769, London, SW1P 9ER. ISD is registered in England with company registration number 06581421 and registered charity number 1141069. All Rights Reserved. Any copying, reproduction or exploitation of the whole or any part of this website without prior written approval from ISD is prohibited.